

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
30 mai 2002 (30.05.2002)

PCT

(10) Numéro de publication internationale
WO 02/43051 A1

(51) Classification internationale des brevets⁷ :

G10L 19/00, H04L 1/20, 12/56

(71) Déposant (pour tous les États désignés sauf US) :
FRANCE TELECOM [FR/FR]; 6, place d'Alleray,
F-75015 Paris (FR).

(21) Numéro de la demande internationale :

PCT/FR01/03671

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : LE SAOUT,
Jean-Yves [FR/FR]; 23, rue des Plages, F-22660 Treleven
(FR). BERNEX, Elodie [FR/FR]; Rubudaz, F-22140 Ton-
quedec (FR). ESTOREZ, Dimitri [FR/FR]; 23 bis, rue de
la Queue en Brie, F-94370 Sucy en Brie (FR).

(22) Date de dépôt international :

21 novembre 2001 (21.11.2001)

(25) Langue de dépôt :

français

(74) Mandataires : LOISEL, Bertrand etc.; Cabinet Plasser-
aud, 84, rue d'Amsterdam, F-75440 Paris Cedex 09 (FR).

(26) Langue de publication :

français

(30) Données relatives à la priorité :

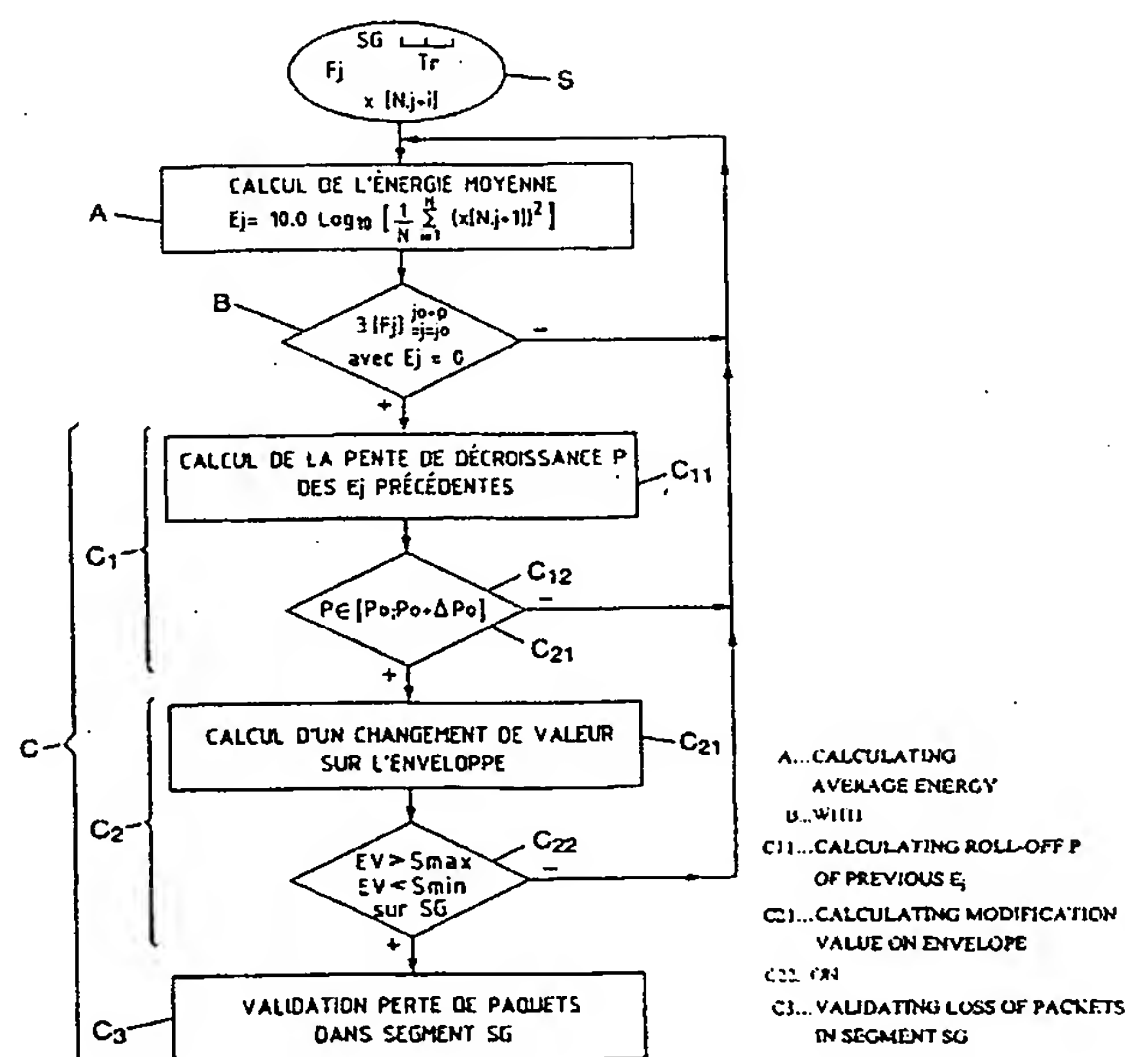
00/15128 23 novembre 2000 (23.11.2000) FR

(81) États désignés (national) : AE, AG, AL, AM, AT, AU, AZ,
BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ,
DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM,

[Suite sur la page suivante]

(54) Title: NON-INTRUSIVE DETECTION OF DEFECTS IN A PACKET-TRANSMITTED SPEECH SIGNAL

(54) Titre : DETECTION NON INTRUSIVE DES DÉFAUTS D'UN SIGNAL DE PAROLE TRANSMIS PAR PAQUETS



(57) Abstract: The invention concerns a method and a system for detecting defects in a packet-transmitted speech signal. The method consists in calculating from a restored speech signal, analysed on at least one analysis window (F_j), the average energy (E_j) of the restored speech signal and in searching (B) the segments of the restored speech signal at null average energy; in validating (C), on at least an existing null average energy window, the loss of packets by frame substitution with silent frames conditional to the existence of a frame substitution. The invention is particularly applicable to IP telephony.

(57) Abrégé : L'invention concerne un procédé et un système de détection des défauts d'un signal de parole transmis par paquets. A partir d'un signal de parole reconstitué, analysé sur au moins une fenêtre d'analyse (F_j), l'énergie moyenne (E_j) du signal de parole reconstitué est calculée (A) et on recherche (B) les segments

[Suite sur la page suivante]

WO 02/43051 A1



HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

- (84) États désignés (*régional*) : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Déclaration en vertu de la règle 4.17 :

— *relative à la qualité d'inventeur (règle 4.17.iv)) pour US seulement*

Publiée :

— *avec rapport de recherche internationale*

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

DETECTION NON INTRUSIVE DES DEFAUTS D'UN
SIGNAL DE PAROLE TRANSMIS PAR PAQUETS.

5 L'invention concerne un procédé et un système de
détection non intrusive des défauts d'un signal de parole
transmis en téléphonie sur réseau de transmission par
paquets.

10 Les possibilités accrues de joindre et de
communiquer avec n'importe quel interlocuteur en un lieu
sensiblement quelconque du globe terrestre, à un instant
quelconque et pour un coût dérisoire, celui d'une
communication locale, par l'intermédiaire du réseau
d'interconnexion mondial de l'Internet, ont suscité un
15 engouement pour la téléphonie sur réseau de transmission
par paquets, plus particulièrement pour la téléphonie sur
IP (Internet Protocol).

20 Contrairement à la téléphonie classique sur réseau
fixe par l'intermédiaire du réseau téléphonique commuté,
la communication en téléphonie sur réseau de transmission
par paquets, en particulier la téléphonie sur IP, présente
le plus souvent une qualité médiocre en l'absence de
garantie d'un niveau de qualité minimum.

25 Un tel défaut de garantie de qualité vocale en
téléphonie sur IP est inhérent à la conception même du
réseau de transmission, lequel favorise l'interconnexion
des interlocuteurs au détriment de la bande passante
finalement allouée au signal de parole, seul un critère de
meilleur effort (best effort) en matière de bande passante
30 allouée étant admissible.

En particulier, le débit des informations transmises, sous forme de paquets numériques représentatifs du signal de parole, diminue au fur et à mesure que les usagers du réseau IP se connectent et utilisent le réseau en téléphonie sur IP ou non.

Le déploiement dans des conditions satisfaisantes des services de téléphonie sur IP nécessite donc la mise en œuvre d'un contrôle de la qualité des services proposés et en particulier la mise en œuvre d'outils de mesure de cette qualité.

La notion de qualité de transmission en téléphonie sur réseau de transmission par paquets repose sur le processus de transmission du signal de parole sous forme de paquets numériques représentatifs de l'information. L'information à transmettre est, en effet, découpée en segments de longueur sensiblement identique, pour une communication donnée, et ainsi transmise séquentiellement.

Par exemple, en téléphonie sur IP, les mécanismes de routage des paquets transmis sur le réseau Internet font que les paquets associés à une même communication téléphonique peuvent, le plus souvent, emprunter des chemins différents. En conséquence, la durée d'acheminement de chaque paquet est variable, car cette durée dépend du chemin effectivement suivi.

Afin de pouvoir reconstituer l'information à la réception, sur un terminal distant, chaque paquet est horodaté à l'émission selon un protocole RTP, pour Real Time Protocol. Un programme applicatif implanté sur le terminal distant reconstitue l'information à la réception à partir des informations d'horodatage introduites selon le protocole RTP.

En outre, afin de ne pas encombrer le réseau Internet, à chaque paquet est attribué une durée de vie, lors de l'émission. A chaque nœud du réseau Internet tel que passerelle ou routeur, cette durée de vie est
5 décrémentée. Lorsque la durée de vie décrémentée atteint la valeur zéro, le paquet correspondant est éliminé.

Les principaux défauts d'une transmission en téléphonie sur IP, encore appelée VoIP, peuvent être mentionnés ci-après :

- 10 - les retards importants, liés aux délais d'acheminement et aux temps de traitement par les équipements.
Ce défaut peut gêner l'interactivité et donc rendre la conversation entre appelant et appelé difficile, voire impossible ;
- 15 - la gigue dans la durée d'acheminement des paquets, ce phénomène se traduisant par des intervalles de temps de longueur différente entre l'instant d'arrivée des paquets.
Un tel phénomène se traduit par un temps de
20 transmission non uniforme ;
- les pertes de paquets, soit parce que ces paquets ont été éliminés lors de l'acheminement, leur durée de vie étant épuisée, suite à l'encombrement des routeurs, soit parce qu'ils sont arrivés au terminal distant avec
25 un retard trop important, ces paquets étant alors détruits à l'arrivée ;
- l'écho lié principalement aux retards élevés et extrêmement variables ;
- la distorsion due au codage du signal de parole en
30 paquets numériques à débit réduit, généralement utilisé en VoIP.

Les défauts précités ont un impact réel sur la qualité du signal de parole reçu et il apparaît très important d'estimer cette qualité, tant au niveau objectif, par mesure des paramètres physiques et acoustiques de ce signal, que subjectif, à partir de tests d'écoute permettant d'identifier ce qui est effectivement perçu par les usagers.

Les outils actuellement disponibles pour assurer une telle estimation peuvent être classés en deux catégories :

- les outils fonctionnant de bout en bout, ces outils suivant la recommandation P 861 de l'UIT-T, série P : Qualité de la transmission téléphonique. Méthodes d'évaluation objective et subjective de la qualité. Mesure objective de la qualité des codecs vocaux fonctionnant en bande téléphonique (300-3400 Hz) ;
- les outils fonctionnant sans intrusion, ces outils suivant la recommandation P 561 de l'UIT-T; série P : Qualité de la transmission téléphonique. Appareils de mesures objectives. Dispositif de mesure en service et sans intrusion. Mesure pour les services vocaux. Chapitres 6 et 7.

Les outils de mesure de bout en bout mettent en œuvre des modèles psycho-acoustiques. Ils injectent un corpus de parole à une extrémité et enregistrent le corpus transmis à l'autre extrémité. Le modèle psycho-acoustique détermine une note de qualité à partir de la comparaison entre les caractéristiques temporelles et fréquentielles du signal source et du signal transmis dégradé. Ce type d'outil nécessite toutefois d'avoir accès aux deux

extrémités de la liaison téléphonique et ne peut être utilisé que dans des cas spécifiques, limités.

Les outils de mesure sans intrusion peuvent, par contre, être connectés en un point quelconque du réseau. Ils fonctionnent au niveau protocolaire, par exemple analyseurs de protocole, ou directement au niveau des données transmises, outils dits I.N.M.D. Les analyseurs de protocole fournissent des statistiques sur les types de protocole mis en œuvre sur le réseau, et, en conséquence, une information sur le type de données véhiculées, courrier électronique, données, voix, les retards, la gigue, les pourcentages de paquets perdus. Ces informations sont obtenues par un décodage des en-têtes des paquets ou directement par l'analyse du signal numérique lorsque l'outil est connecté à un réseau commuté. Dans le cas des outils I.N.M.D, la qualité des communications est appréciée par l'intermédiaire de la mesure des paramètres de transmission, niveaux de signal et de bruit, retard, affaiblissement d'écho.

Les outils de mesure de bout en bout présentent un intérêt, car ils permettent de qualifier sur une échelle à cinq niveaux la qualité d'un réseau, d'un équipement ou d'un ensemble d'équipements. Ce type d'outil implique par contre l'obligation de disposer de deux accès à la ligne de transmission, l'un en amont et l'autre en aval du système à qualifier. L'un des accès sert à l'injection d'un corpus stimulus de mesure, l'autre à l'enregistrement du corpus dégradé.

Ce mode opératoire devient un inconvénient rédhibitoire, lorsqu'un des accès n'est pas accessible, soit par que géographiquement éloigné, soit par ce que, dans le cadre

de l'interconnexion des réseau, il n'appartient pas à l'opérateur qui conduit les tests. Il ne peut être mis en œuvre dans le cadre de la téléphonie sur réseau de transmission par paquets, sauf, le cas échéant, dans des cas très limités.

Les outils sans intrusion fonctionnent à partir d'un seul accès et ne nécessitent l'injection d'aucun corpus stimulus. Toutefois, le type des résultats fournis par ce type d'outil, tels que des statistiques sur des paramètres de transmission, caractérise le fonctionnement du système support de cette transmission plutôt que la qualité du signal de parole transmis. Il est alors nécessaire d'utiliser des modèles de transfert, tel que le modèle E de la recommandation G 107 de l'UIT-T, afin de pouvoir disposer d'une estimation de la qualité perçue à partir des caractéristiques de transmission.

La présente invention a pour objet de remédier aux insuffisances et limitations des outils de l'art antérieur, notamment dans le cadre de la téléphonie sur réseau de transmission par paquets.

En particulier, la présente invention a pour objet la mise en œuvre d'un procédé et d'un système de détection non intrusive des défauts d'un signal de parole transmis en téléphonie sur réseau de transmission par paquets permettant de détecter les défauts perceptibles par les usagers directement sur le signal de parole transmis.

Un autre objet de la présente invention est en outre la mise en œuvre d'un procédé et d'un système de détection non intrusive des défauts d'un signal de parole transmis en téléphonie sur réseau de transmission par paquets permettant de caractériser la qualité perçue du

signal de parole à partir d'une bibliothèque de dégradations, cette bibliothèque caractérisant les défauts, tant d'un point de vue objectif par l'énergie, le spectre, le *pitch* (fréquence fondamentale) du signal de parole, que d'un point de vue subjectif par une définition des types de défauts, qualification verbale de ces défauts et quantification de la gêne provoquée par chaque type de défaut.

Un autre objet de la présente invention est en outre la mise en œuvre d'un procédé et d'un système de détection non intrusive des défauts d'un signal de parole transmis en téléphonie sur réseau de transmission par paquets permettant de déterminer une qualité perçue par l'utilisateur.

Un autre objet de la présente invention est, également, la mise en œuvre d'un système de détection non intrusive des défauts d'un signal de parole transmis, permettant, d'une part, de désencapsuler le signal de parole contenu dans les paquets, afin de reconstituer l'information vocale, lorsque ce système est connecté sur une portion de réseau de transmission par paquets, et, d'autre part, de récupérer directement le signal de parole lorsque ce système est connecté sur une portion du réseau RTC.

Un autre objet de la présente invention est également, à partir d'un signal de parole reconstitué, la mise en œuvre d'un procédé et d'un système de détection non intrusive des défauts d'un signal de parole transmis en téléphonie sur réseau de transmission par paquets ou RTC permettant de détecter les différents types de dégradation présents sur ce signal de parole, leur durée

et leur occurrence, et leur influence sur la qualité vocale perçue.

Un autre objet de la présente invention est, enfin, la mise en œuvre d'un système de détection non intrusive des défauts d'un signal de parole transmis en téléphonie sur IP permettant de cumuler les avantages des analyseurs sans intrusion et les outils de mesure de bout en bout de l'art antérieur, en l'absence de leurs inconvénients spécifiques.

Le procédé de détection des défauts objectifs d'un signal de parole transmis par paquets en téléphonie sur réseau de transmission par paquets, objet de la présente invention, est remarquable en ce qu'il consiste, à partir d'échantillons d'un signal de parole reconstitué représentatif de ce signal de parole transmis et analysés sur au moins une fenêtre d'analyse:

- à calculer l'énergie moyenne du signal de parole reconstitué dans cette fenêtre d'analyse,
- à rechercher, pour une succession de fenêtres d'analyse, les segments du signal de parole reconstitué à énergie moyenne nulle, et, sur l'existence d'au moins une fenêtre d'analyse d'énergie moyenne nulle,
- à valider la perte de paquets avec substitution par des trames de silence conditionnellement à l'existence d'une substitution de trame.

Le système de détection non intrusive des défauts d'un signal de parole transmis en téléphonie sur réseau de transmission par paquets, objet de l'invention, est remarquable en ce qu'il comporte au moins :

- un module de capture et de reconstitution du signal de parole reconstitué ;

- un module de base de données, comportant une bibliothèque de signaux de défauts objectifs, liés aux caractéristiques physiques du signal de parole, et de défauts subjectifs, liés au typage de la gêne occasionnée à des utilisateurs par les signaux de défauts objectifs, cette bibliothèque étant construite à partir d'une analyse statistique du type de dégradation ou de défaut présent dans le signal de parole reconstitué ;
- 10 - un module de détection de défauts présents sur le signal de parole reconstitué et en conséquence sur le signal de parole transmis, et
- un module de calcul et d'affichage par comparaison des défauts objectifs présents sur le signal de parole reconstitué aux signaux de défauts objectifs et/ou
- 15 signaux de défauts subjectifs de la bibliothèque, permettant, à partir d'une analyse statistique du type de dégradation ou de défaut présent dans le signal de parole reconstitué et dans le signal de parole transmis, d'afficher des paramètres de qualité de ce
- 20 signal de parole transmis.

Le procédé et le système objets de l'invention trouvent application à l'amélioration de la qualité des communications téléphoniques en téléphonie sur réseau de

25 transmission par paquets, notamment en téléphonie sur IP.

Ils seront mieux compris à la lecture de la description et à l'observation des dessins dans lesquels :

- la figure 1 représente, à titre illustratif, un organigramme des étapes essentielles permettant la mise
- 30 en œuvre du procédé objet de la présente invention ;

- la figure 2a représente un chronogramme d'un échantillon de signal de parole et de l'enveloppe de ce signal de parole, après décodage, lorsqu'un tel signal, transmis en téléphonie sur IP, est affecté d'une perte de cinq paquets ;
- la figure 2b représente un chronogramme d'un échantillon de signal de parole, après décodage de l'enveloppe de ce signal de parole et de la dérivée de cette enveloppe, lorsqu'un tel signal, transmis en téléphonie sur IP, est affecté d'une double perte de cinq paquets ;
- la figure 2c représente un chronogramme des valeurs de la dérivée de l'enveloppe d'un signal de parole transmis en téléphonie sur IP en fonction de l'emplacement du défaut, perte de paquets, dans l'échantillon de parole considéré ;
- la figure 2d représente un chronogramme d'un échantillon de signal de parole dans lequel le changement de valeur d'énergie aux extrémités d'un changement de la valeur de l'enveloppe de ce signal de parole permet de valider la perte de paquets dans un segment d'activité vocale ;
- la figure 3a représente, à titre illustratif, un schéma synoptique d'un système de détection non intrusive des défauts d'un signal de parole transmis en téléphonie sur réseau de transmission par paquets, conforme à l'objet de la présente invention ;
- la figure 3b représente, à titre illustratif, un détail de réalisation d'un module de capture et de reconstitution d'un signal de parole transmis en téléphonie sur réseau de transmission par paquets

incorporé dans le système objet de l'invention représenté en figure 3a et permettant d'obtenir un signal de parole reconstitué, que le signal de parole transmis en téléphonie sur réseau de transmission par paquets soit analysé au niveau d'un tel réseau ou au contraire sur un réseau RTC ;

- la figure 3c représente, à titre illustratif, une architecture de modules logiciels spécifiques implantés au niveau de l'unité centrale du système objet de l'invention, et permettant la mise en œuvre de l'ensemble des fonctionnalités du système objet de la présente invention.

Une description plus détaillée du procédé et du système de détection non intrusive des défauts d'un signal de parole transmis en téléphonie sur réseau de transmission par paquets, conformes à l'objet de la présente invention, sera maintenant donnée en liaison avec la figure 1 et des figures suivantes.

D'une manière générale, on indique que le procédé objet de la présente invention, est mis en œuvre à partir d'un signal de parole extrait du réseau de transmission de ce signal de parole, ce dernier étant transmis en téléphonique sur réseau de transmission par paquets.

De manière plus spécifique, on indique que le procédé, objet de la présente invention, concerne tout signal de parole transmis sous forme de paquets, ce signal de parole et les paquets correspondants, supports de ce dernier, pouvant être détectés soit sur un réseau de transmission de type réseau local, encore désigné par réseau LAN, pour Local Area Network, ou sur un réseau longue distance, encore désigné réseau WAN, pour Wide Area

Network, ou tout autre réseau de transmission par paquets, -ATM (Asynchronous Transmission Mode), FR (Frame Relay) notamment-, ou, le cas échéant, sur le réseau téléphonique commuté RTC, ainsi qu'il sera décrit ultérieurement dans la description. En référence à la figure 1, le procédé
 5 objet de la présente invention consiste, à partir d'échantillons d'un signal de parole reconstitué, représentatif bien entendu du signal de parole transmis, à effectuer une analyse du signal de parole reconstitué sur
 10 au moins une fenêtre d'analyse de durée déterminée.

En référence à la figure 1, on considère une étape de départ S, dans laquelle on dispose sur la fenêtre d'analyse F_j , où j désigne l'indice de la fenêtre d'analyse précitée, d'un nombre N d'échantillons de ce
 15 signal de parole reconstitué, chaque échantillon étant noté $x[N.j+i]$, où j désigne l'indice de la fenêtre d'analyse et i l'indice de l'échantillon dans la fenêtre de rang j .

A partir de l'analyse conduite sur les échantillons du signal de parole reconstitué précité, le
 20 procédé, objet de la présente invention, consiste, en une étape A, à calculer l'énergie moyenne du signal de parole reconstitué dans la fenêtre d'analyse F_j précitée. La valeur de l'énergie moyenne du signal de parole
 25 reconstitué vérifie la relation (1) :

Relation 1

$$E_j = 10.0 \text{Log}_{10} \left[\frac{1}{N} \sum_{i=1}^N (x[N.j+i])^2 \right]$$

L'étape A précitée est suivie d'une étape B consistant à rechercher, pour une succession de fenêtres d'analyse successives, les segments du signal de parole reconstitué dont l'énergie moyenne est sensiblement nulle.

5 L'étape B est représentée par une étape de test, le test précité consistant à vérifier l'existence des segments du signal de parole reconstitué précité satisfaisant à la relation (2) :

10 Relation 2

$$\exists [F_j]_{j=j_0}^{j_0+p} \text{ avec } E_j = 0$$

On comprend en particulier que les segments du signal de parole peuvent être constitués par une pluralité de
15 fenêtres d'analyse successives.

L'étape B précitée est alors suivie d'une étape C consistant à valider la perte de paquets avec substitution par des trames de silence conditionnellement à l'existence
20 d'une substitution de trames.

On comprend bien sûr que l'opération de substitution de paquets perdus par des trames de silence, effectuée directement par le processus de codage et de transmission du signal de parole, constitue l'un des
25 défauts objectifs permettant de qualifier le niveau de qualité du signal de parole transmis à partir du signal de parole reconstitué, conformément au procédé objet de la présente invention.

Un mode de mise en œuvre plus spécifique de
30 l'étape C consistant à valider la perte des paquets avec

substitution par des trames de silence conditionnellement à l'existence d'une substitution de trames précédente, sera maintenant décrit ci-après.

5 Selon un aspect avantageux de mise en œuvre du procédé objet de la présente invention, outre l'étape de calcul de l'énergie moyenne du signal de parole reconstitué dans chaque fenêtre d'analyse, et la détection de segments du signal de parole reconstitué pour lesquels l'énergie moyenne est nulle au moyen du test réalisé à
10 l'étape B, un critère supplémentaire C_1 permet de valider la présence d'une perte de paquets substitués par du silence après plusieurs substitutions, par répétition ou interpolation avec baisse du niveau d'énergie lorsque, par exemple, une telle opération est réalisée lors du
15 processus de décodage du signal de parole.

Dans ces conditions, l'étape consistant à valider la perte de paquets peut comporter une étape C_{11} consistant à calculer la pente de décroissance de l'énergie d'excitation de trames antérieures successives
20 sur les fenêtres de rang $j-3$, $j-2$, $j-1$, les énergies moyennes correspondantes étant notées E_{j-3} , E_{j-2} et E_{j-1} respectivement. En effet, c'est la décroissance de l'énergie des trames ou fenêtres substituées qui précède les trames à énergie nulle. De façon générale, si la trame
25 de rang k est la première trame à énergie nulle détectée, alors, il existe une décroissance d'énergie entre les trames successives de rang $k-n$ et $k-(n-1)$, puis successivement $k-(n-1)$ et $k-(n-2)$ et ainsi de suite, où n désigne le nombre de paquets substitués avec atténuation
30 d'énergie avant la substitution par du silence. C'est le cas du décodeur G 723.1.

L'étape C_{11} est ainsi suivie d'une étape C_{12}) consistant à valider la pente de décroissance de l'énergie d'excitation déterminée à partir du signal de parole reconstitué par rapport à la pente de décroissance théorique caractéristique du décodeur utilisé.

Dans ces conditions, l'étape C_{12}) consiste alors, ainsi que représenté sur la figure 1, à effectuer un test sur la valeur de la décroissance de pente P précitée, en une étape de test C_{12}). L'étape de test précitée, sur la valeur de pente P , consiste à vérifier l'appartenance de la valeur de la pente P calculée pour chaque valeur successive à une plage de valeur de pente selon la relation (3) :

Relation (3)

$$P \in [P_0; P_0 + \Delta P_0]$$

Sur réponse négative au test C_{12}) précité, un retour au calcul de l'énergie moyenne du signal de parole reconstitué pour chaque fenêtre d'analyse à l'étape A) est effectué.

Sur réponse positive au test C_{12}), l'étape suivante C_2) consistant à valider la perte de paquets en fonction de l'existence d'un défaut objectif dans un segment d'activité vocale peut alors consister à calculer, dans une étape C_{21}), une valeur de changement de la valeur de l'enveloppe du signal de parole reconstitué, ce changement de valeur de l'enveloppe pouvant consister en un calcul de l'amplitude de la valeur d'enveloppe EV, ainsi qu'il sera décrit ultérieurement dans la description. L'étape de calcul C_{21}) est alors suivie d'une étape de test C_{22})

consistant à comparer la valeur de changement EV précitée à au moins une valeur de seuil. Sur la figure 1, la valeur de changement d'enveloppe EV est comparée à une valeur de seuil maximale S_{\max} , respectivement à une valeur de seuil minimale S_{\min} par comparaison de supériorité $EV > S_{\max}$, respectivement d'infériorité $EV < S_{\min}$ sur le segment de parole SG considéré.

Sur réponse négative au test C_{22}), un retour à l'étape A de calcul de l'énergie moyenne sur chaque fenêtre d'analyse est réalisé. Au contraire, une réponse positive à l'étape de test C_{21}) révèle l'existence d'une perte de paquets dans un segment d'activité vocale SG engendrant un défaut objectif et, en conséquence, un défaut subjectif, alors que la réponse négative à la comparaison C_{22}) révèle l'absence de défaut subjectif.

Deux modes de réalisation particuliers non limitatifs de l'étape de calcul C_{21}) d'un changement de valeur sur l'enveloppe du signal de parole reconstitué seront maintenant donnés en liaison avec les figures 2a à 2d.

La figure 2a représente successivement un échantillon du signal de parole reconstitué et la valeur de l'énergie moyenne dans une fenêtre d'analyse F_j , ces chronogrammes étant représentés en niveaux d'énergie en décibels en ordonnées, et en durée d'échantillon en secondes en abscisses.

L'échantillon du signal de parole reconstitué $x(t)$ correspond à un signal de parole codé selon la norme G 723.1 et dégradé par la perte de cinq paquets, puis traité dans un décodeur G 723.1. On rappelle en particulier qu'une stratégie de masquage des erreurs dues

à des effacements de trame, c'est-à-dire de perte de paquets, peut être intégrée dans les décodeurs et, en particulier, dans les décodeurs de type G 723.1. Si un effacement de trame se produit, un algorithme de substitution remplace la trame perdue par une trame de silence, par du bruit ou par répétition, ou interpolation, en fonction du type de codage utilisé et en fonction des caractéristiques de la dernière trame reçue, activité ou absence d'activité de parole, trame voisée ou non voisée. Ainsi, en fonction du type de processus de codage/décodage utilisé en fonction du nombre de paquets perdus et en fonction de la position de ces paquets perdus dans le signal de parole, les défauts sont perçus de manière différente par les usagers.

Sur la figure 2a, on peut observer que les deux premières trames effacées ont été substituées par interpolation et diminution du niveau d'énergie du signal de parole reconstitué, mais que les trois trames suivantes n'ont fait l'objet d'aucune interpolation et correspondent à un signal sensiblement nul. On remarque en outre que l'énergie est nulle durant un intervalle de temps correspondant à trois paquets.

La mise en œuvre de l'étape de calcul d'un changement de valeur sur l'enveloppe du signal de parole reconstitué, à l'étape C₂₁) et le test C₂₂), permet en fait de localiser l'emplacement du défaut dans le corpus vocal, perte de paquets dans un segment de parole ou dans un segment de silence.

Lorsque le défaut apparaît dans l'activité vocale proprement dite, ce défaut est pris en compte, car, dans ces conditions, ce défaut objectif provoque chez

l'utilisateur une notion de défaut subjectif qui est perçue comme telle par ce dernier. Lorsqu'au contraire, le défaut objectif apparaît en dehors de l'activité vocale, ce dernier n'a pas d'incidence sur la qualité vocale perçue et il n'est donc pas perçu comme un défaut subjectif.

Des phénomènes de réverbération et la production de la parole par modulation de l'excitation par les conduits vocaux ont pour conséquence que l'arrêt de l'activité vocale ou la reprise de cette dernière ne se font jamais de façon brutale. Ainsi, l'amplitude du signal de parole, et bien entendu du signal de parole reconstitué, ne présente donc que des variations sensiblement continues. La présence d'un défaut dans un échantillon de parole, tel que la perte d'un paquet IP dans un segment d'activité vocale, engendre donc automatiquement une rupture dans l'amplitude du signal de parole ou du signal de parole reconstitué.

La mise en œuvre des étapes C_{21}) et C_{22}) a pour objet de localiser la discontinuité ainsi engendrée.

Selon un premier mode de réalisation non limitatif de l'étape C_{21}) de calcul d'un changement de valeur sur l'enveloppe, ce changement, c'est-à-dire cette discontinuité, peut être localisé par le calcul de la dérivée de l'enveloppe du signal de parole ou du signal de parole reconstitué.

Dans ce but, le signal de parole reconstitué $x(t)$, c'est-à-dire les échantillons $x[N.j+i]$ peuvent être soumis à un filtrage passe-bas délivrant l'enveloppe du signal de parole reconstitué $y(t)$ correspondant sensiblement à l'amplitude du signal de parole reconstitué.

La valeur de la dérivée du signal de parole devient alors importante, lorsque l'amplitude du signal de parole présente une discontinuité. C'est en particulier le cas aux extrémités d'un segment de parole correspondant à un défaut avec substitution par du silence.

En ce qui concerne le calcul de l'amplitude du signal de parole reconstitué, et finalement de l'enveloppe de ce dernier, ce calcul peut être effectué par segments. Ainsi, une valeur de l'amplitude et de l'enveloppe précitée peut être déterminée pour chaque fenêtre d'analyse. L'énergie moyenne de chaque trame est en conséquence une représentation de l'enveloppe du signal de parole reconstitué et donc du signal de parole.

Sur la figure 2b, on a représenté successivement un échantillon du signal de parole reconstitué, l'énergie moyenne dans la fenêtre d'analyse exprimée en dB et, enfin, la dérivée de l'enveloppe du signal de parole reconstitué, également exprimée en dB. L'échantillon du signal de parole correspond à un échantillon codé selon le processus de codage/décodage G 723.1 dégradé par deux pertes de cinq paquets, puis traité par un décodeur G 723.1. La première perte de paquets IP intervient dans un segment d'activité vocale, tandis que la deuxième perte de paquets IP intervient dans un segment de non-activité vocale. De la même manière que dans le cas de la figure 2a, on constate que les deux premières trames effacées ont été substituées par interpolation de la dernière trame reçue et que les trois trames suivantes n'ont fait l'objet d'aucune interpolation (elles ont été substituées par du silence), et cela pour les deux défauts.

On note également que l'énergie de l'échantillon du signal de parole, échantillon vocal, est nulle durant un intervalle de temps correspondant à trois paquets.

En outre, on constate également que le module de la dérivée de l'enveloppe prend des valeurs importantes lors du passage de la deuxième à la troisième trame du défaut, ainsi qu'au passage de la dernière trame du défaut à la nouvelle trame correctement reçue.

Des investigations par des tests de la détection de défauts simulés ont montré qu'il existe une valeur de seuil S au-dessus de laquelle le module de la dérivée de l'enveloppe du signal de parole reconstitué valide en fait le critère d'une perte de paquets avec substitution par des trames de silence, conditionnellement à l'existence d'une substitution de trame.

Le calcul de la dérivée de l'enveloppe du signal de parole, ou du signal de parole reconstitué, permet, en référence à la figure 2c, d'identifier l'emplacement du défaut objectif dans le corpus vocal. En effet, en référence à la figure 2c précitée, on constate que les valeurs de la dérivée aux extrémités du défaut sont plus faibles lorsque la dégradation intervient en l'absence d'activité vocale, c'est-à-dire dans une zone de silence. On peut ainsi constater, au niveau de la figure 2c précitée, l'existence d'un écart de 30 à 40 dB sur la valeur de la dérivée en fonction de la position du défaut, lorsque ce défaut correspond à un segment d'activité vocale ou, au contraire, à un segment exempt d'activité vocale.

Ainsi, lorsque, en référence à la figure 2c, le défaut introduit dans l'échantillon de parole est une

perte de cinq paquets, les valeurs de la dérivée de l'enveloppe du signal de parole correspondent sur la figure précitée à la transition de fin du défaut, nouvelle trame correctement reçue. La différence d'amplitude au
5 niveau du module de la dérivée du signal de parole ou du signal de parole reconstitué en fonction de l'emplacement du défaut objectif dans l'échantillon de parole considéré, met ainsi en évidence la possibilité de détecter dans
10 transitions, dernière trame reçue - apparition du défaut et apparition du défaut - nouvelle trame reçue.

Toutefois, la valeur de seuil S liée à la méthode d'extraction de l'enveloppe du signal de parole doit en fait être calibrée en fonction des équivalents pour la
15 sonie à l'émission.

On rappelle que la sonie à l'émission désigne l'efficacité du passage de l'onde acoustique au signal de parole électrique.

Afin de se démarquer de la dépendance par rapport
20 aux équivalents pour la sonie des terminaux à l'émission, un deuxième mode de mise en œuvre de l'étape de calcul de changement de valeur sur l'enveloppe C_{21}) et du test C_{22}) peut être proposé, ainsi qu'il sera décrit en liaison avec la figure 2d.

25 Ainsi que représenté sur la figure précitée, le signal de parole ou signal de parole reconstitué et, en particulier, le signal d'enveloppe exprimé en dB en fonction du temps, présente une variation d'énergie significative aux extrémités de chaque trame substituée
30 par du silence dans la zone d'activité vocale. Cette variation d'énergie aux extrémités permet alors de

s'affranchir de la dépendance du réglage du terminal émetteur, c'est-à-dire de la dépendance à l'efficacité du transducteur d'émission et aux réglages audio du terminal utilisé, ainsi que précédemment mentionné dans la description relativement à la méthode consistant à
5 calculer la dérivée de l'enveloppe du signal de parole.

Dans le mode de réalisation considéré correspondant à la figure 2d, l'utilisation d'un algorithme de détection d'activité vocale permet de
10 s'affranchir du réglage du terminal d'émission. Ce type d'algorithme détermine un seuil d'amplitude ou d'énergie permettant de tronçonner l'échantillon de parole en segments d'activité vocale et en segments de non-activité vocale. La détermination de ce seuil est indépendante du
15 réglage du terminal d'émission. On discrimine les défauts situés dans des tronçons d'activité vocale, en comparant l'énergie de la dernière trame reçue avant le défaut et l'énergie de la première trame reçue après le défaut, au seuil déterminé par l'algorithme précité. Si les énergies
20 sont supérieures au seuil, alors le défaut s'est produit en pleine activité vocale. Par contre, si les énergies sont inférieures au seuil, alors le défaut s'est produit en l'absence d'activité vocale.

Enfin, l'étape C₃) de validation de perte de
25 paquets dans un segment de parole permet ensuite d'attribuer une note de qualité à la communication suivant l'existence d'un tel défaut.

Une description plus détaillée d'un système de
détection non intrusive des défauts d'un signal de parole
30 transmis en téléphonie sur réseau de transmission par

paquets, conforme à l'objet de la présente invention, sera maintenant donnée en liaison avec les figures 3a à 3c.

5 D'une manière générale, on rappelle, en référence à la figure 3a, que le système objet de la présente invention est adapté aux fins de fonctionner directement à partir du signal de parole, de manière à déterminer une qualité perçue par l'utilisateur. Ainsi, le système précité met en œuvre un dispositif permettant de désencapsuler le signal de parole contenu dans les paquets
10 afin de reconstituer l'information vocale lorsque le système objet de l'invention est connecté sur un réseau de transmission par paquets, ou le cas échéant, de récupérer directement le signal vocal lorsque le système objet de l'invention est connecté sur une portion de transmission
15 du réseau téléphonique commuté.

Une fois que le signal de parole est reconstitué, selon un signal de parole reconstitué représentatif de ce signal de parole, un traitement de signal approprié détecte les différents types de dégradation présents dans
20 le signal de parole considéré, ainsi que leur durée et leur occurrence. Ces défauts sont ensuite reliés à une influence sur la qualité vocale perçue dans les conditions qui seront explicitées ci-après.

En référence à la figure 3a, le système de
25 détection, objet de la présente invention, comporte un module 1 de capture et de reconstitution du signal de parole transmis par paquets, ce module 1 délivrant un signal de parole reconstitué, noté spr.

En outre, ainsi que représenté sur la figure 3a
30 précitée, le système objet de l'invention comprend une base de données, portant la référence 2 et représentée

symboliquement par un disque dur HDD, cette base de données comportant une bibliothèque de signaux de défauts objectifs liés aux caractéristiques physiques du signal de parole, et bien entendu du signal de parole reconstitué spr, et de défauts subjectifs liés au typage de la gêne occasionnée à des utilisateurs par les signaux de défauts objectifs.

D'une manière plus spécifique, on indique que la bibliothèque est construite à partir d'une analyse statistique du type de dégradation ou de défaut présent dans le signal de parole reconstitué spr dans les conditions qui seront explicitées de manière plus détaillée ultérieurement dans la description.

En outre, ainsi que représenté sur la figure 3a précitée, le système objet de l'invention comprend un module 3 de détection des défauts présents sur le signal de parole reconstitué spr et, en conséquence, sur le signal de parole transmis sp. Sur la figure 3a précitée, les modules constitutifs du système de détection non intrusive, objet de la présente invention, constitués essentiellement par des modules logiciels, sont représentés de manière symbolique par un ovale, afin de distinguer ces modules des modules matériels constitutifs du système objet de la présente invention.

En outre, un module 3 de calcul et d'affichage par comparaison des défauts objectifs présents sur le signal de parole reconstitué spr aux signaux de défauts objectifs et/ou aux signaux de défauts subjectifs contenus dans la bibliothèque des défauts, permet, à partir d'une analyse statistique du type de dégradation ou de défaut présent dans le signal de parole reconstitué spr et dans le signal

de parole transmis sp, d'afficher des paramètres de qualité du signal de parole transmis sp.

En référence à la figure 3a, on indique que, après détection des types de défauts par le module 3 de
5 détection de défauts présents sur le signal de parole reconstitué spr, on dispose en fait de variables de types de défauts, notées tdf, lesquelles sont comparées aux signaux de défauts objectifs et/ou subjectifs de la bibliothèque, ces signaux étant notés tdfb au niveau du
10 module 4 de calcul et d'affichage par comparaison des défauts objectifs.

Ainsi que représenté sur la figure 3a, on comprend bien entendu que pour assurer le calcul et l'affichage par comparaison des défauts objectifs précités, le module 4 de
15 calcul comprend avantageusement un module logiciel 4a assurant le traitement de comparaison entre les valeurs de type de défauts tdf et les signaux de défaut de la bibliothèque tdfb, ce module logiciel 4a étant bien entendu associé à une unité de calcul 4b, comportant une
20 unité centrale UC de traitement et bien entendu une unité d'affichage notée UF.

Ainsi, le système de détection non intrusive de défauts d'un signal de parole, objet de la présente invention, peut être réalisé à partir du module de capture
25 et de reconstitution du signal de parole reconstitué spr, constitué par un module matériel de capture et de reconstitution, lequel peut alors être interconnecté par l'intermédiaire d'une liaison numérique à un micro-ordinateur de type PC par exemple, ou à un calculateur
30 dédié à base de processeurs de traitement de signal DSP.

La réalisation du système objet de la présente invention sur un matériel informatique de type ordinateur PC présente l'intérêt d'une grande souplesse de mise en œuvre, dans la mesure où, bien entendu, les modules 3 et 4 de détection des types de défaut et de calcul par comparaison peuvent être constitués par des programmes directement implantés en mémoire morte, ou non volatile, de l'ordinateur et chargés en mémoire de travail lors du traitement.

10 Une description plus détaillée du module 1 de capture et de reconstitution du signal de parole transmis en un signal de parole reconstitué spr sera maintenant donnée en liaison avec la figure 3b.

D'une manière générale, pour réaliser la mise en œuvre du module 1 de capture et de reconstitution du signal de parole transmis, il est nécessaire de tenir compte du fait que le système objet de la présente invention doit être alimenté par le signal de parole transmis par paquets extraits, soit du réseau de transmission numérique, soit à partir du réseau téléphonique commuté.

Dans le cas d'un réseau local de type LAN, pour *Local Area Network*, ou d'un réseau longue distance de type WAN, pour *Wide Area Network*, ou de tout autre type de réseau de transmission par paquets (ATM, FR, GSM) servant de support à tout signal de parole émis sous forme de paquets, le module 1 de capture et de reconstitution comporte un analyseur de protocole 10, ainsi que représenté sur la figure 3b, permettant de capturer et de désencapsuler la voix pour la restituer sous forme de signal de parole reconstitué spr. Dans ces conditions, la

conversation entre les usagers de deux terminaux peut alors être enregistrée sous forme de fichier, fichier binaire ou fichier son, ou transférée directement au système objet de l'invention.

5 Dans le cas d'une installation du système objet de la présente invention sur une portion de réseau téléphonique commuté, réseau RTC, le signal de parole peut alors être directement accessible par l'intermédiaire d'une sonde 1_1 à haute impédance, connectée au niveau
10 d'une liaison numérique à quatre fils par exemple.

L'extraction du signal de parole, à partir du signal de parole transmis, pour engendrer le signal de parole reconstitué spr, peut alors être réalisée par différents systèmes suivant le type de connexion au
15 réseau, c'est-à-dire la connexion par l'intermédiaire de l'analyseur de protocole 1_0 , ou respectivement la connexion par l'intermédiaire de la sonde à haute impédance 1_1 .

En ce qui concerne le processus de
20 désencapsulation de la voix, ce processus peut être mis en œuvre dans le cas de la connexion de l'analyseur de protocole 1_0 par un programme applicatif de désencapsulation, c'est-à-dire un programme utilisant les mêmes processus de codage/décodage, définis par exemple
25 selon les normes G 723.1, G 729, G 711 et autres, que les terminaux ou passerelles du réseau.

Dans ces conditions, l'échantillon de parole prélevé par l'analyseur de protocole 1_0 est donc sensiblement identique au corpus vocal que restituerait un
30 terminal placé au même endroit que l'analyseur de protocole 1_0 dans le réseau considéré. Dans ces

conditions, l'échantillon ainsi produit par l'analyseur de protocole l_0 comprend bien entendu le signal de parole affecté des défauts dus à la transmission et, en particulier, de la perte des paquets d'information ou à l'introduction de bruit ambiant à l'extrémité d'émission du signal de parole sp.

Dans le cas d'un raccordement de la sonde à haute impédance l_1 au réseau téléphonique commuté RTC, le signal de parole sp est directement accessible sous forme de trames MIC à 64 kbits et l'opération de désencapsulation de la voix n'est alors pas nécessaire. On dispose ainsi du signal de parole reconstitué spr directement à partir de la sortie de la sonde à haute impédance.

Une description plus détaillée du module 2 constitutif de la base de données comportant la bibliothèque de signaux de défauts objectifs sera maintenant donnée ci-après.

D'une manière générale, on indique qu'une classification des dégradations ou défauts présents sur le signal de parole est réalisée en relation avec la perception de ces défauts par un ou plusieurs usagers. On obtient ainsi une bibliothèque de types de dégradations perçues. La bibliothèque des défauts précitée est ainsi obtenue grâce à une étude subjective reposant sur des tests d'écoute par un groupe d'utilisateurs utilisant la méthode de catégorisation libre. Pour obtenir la bibliothèque des défauts précitée, l'on procède à une audition par un ensemble de sujets d'un certain nombre de signaux vocaux, constitués par des échantillons de parole affectés par une perte de paquets et qu'ils doivent

regrouper entre eux lorsque la dégradation leur semble sensiblement identique.

Il est alors demandé aux sujets procédant à l'audition de qualifier, c'est-à-dire de verbaliser les groupes de défauts perçus.

Parmi les défauts perçus par les sujets précités, l'un des défauts qui ressort le plus clairement est celui qui est qualifié de coupure. Il s'agit d'une perte de paquets IP avec substitution par une ou plusieurs trames de silence.

La bibliothèque des défauts a ainsi pour objet d'assurer une classification des différents types de défauts objectifs présents sur le signal de parole reconstitué spr, et donc sur le signal de parole sp, en fonction des types de décodeurs utilisés et bien entendu les conditions de présence ou d'absence, génération discontinue, de trame désignée par DTX, selon la terminologie habituelle. A ces types de défauts objectifs, sont associés des types de défauts subjectifs ainsi que représenté par exemple dans le tableau ci-après :

DEFAUTS SUBJECTIFS

TESTS		GROUPEs	GENE
1	Codeur : G723.1 Débits : 6.3 kbits/s DTX : Présent Voix : Femme	<ul style="list-style-type: none"> ▪ Dégradations non perçues ▪ Coupures ▪ Métalliques : voix de robots, métalliques+résonances et atténuations moins électroniques ▪ Fortes atténuations et résonances ▪ Voix déformées 	K ₁₁ K ₁₂ K ₁₃ K ₁₄ K ₁₅
2	Codeur : G723.1 Débits : 6.3 kbits/s DTX : Présent Voix : Homme	<ul style="list-style-type: none"> ▪ Dégradations non perçues ▪ Coupures et baisse d'énergie ▪ Métalliques et métalliques avec bips ▪ Métalliques (très robot) 	K ₂₁ K ₂₂ K ₂₃ K ₂₄
3	Codeur : G723.1 Débits : 6.3 kbits/s DTX : Absent Voix : Homme	<ul style="list-style-type: none"> ▪ Dégradations non perçues ▪ Bip, aigu, strident ▪ Métalliques, légers grésillements ▪ Métalliques ▪ Sur le mot "planning" ▪ Coupures 	K ₃₁ K ₃₂ K ₃₃ K ₃₄ K ₃₅ K ₃₆
4	Codeur : G723.1 Débits : 6.3 kbits/s DTX : Absent Voix : Femme	<ul style="list-style-type: none"> ▪ Dégradations non perçues ▪ Etouffement du mot, baisse de volume et saut dans la parole ▪ Baisse d'énergie, presque une coupure ▪ Bips, bruits parasites, bruits de fond et une baisse d'énergie ▪ Métallique avec une baisse d'énergie et des sons parasites 	K ₄₁ K ₄₂ K ₄₃ K ₄₄ K ₄₅
5	Codeur : G729 Débits : 8 kbits/s DTX : Absent Voix : Femme	<ul style="list-style-type: none"> ▪ Dégradations non perçues ▪ Voix déformées et baisse d'intensité ▪ Sur le mot "entraîner" (accentuation sur "r" et "tr") ▪ Coupures 	K ₅₁ K ₅₂ K ₅₃ K ₅₄

.../...

TESTS		GROUPE	GENE
6	Codeur : G729 Débits : 8 kbits/s DTX : Absent Voix : Homme	<ul style="list-style-type: none"> ▪ Dégradations non perçues ▪ Légères coupures, voire de fortes atténuations (sur le mot "gare") ▪ Soufflement, voix bruitée (comme si on soufflait dans un micro) ▪ Coupures ▪ Très fortes atténuations, presque des coupures ▪ Coupures sur le mot "perturbation" ▪ Légers effets métalliques avec des atténuations 	K ₆₁ K ₆₂ K ₆₃ K ₆₄ K ₆₅ K ₆₆
7	Codeur : G729 Débits : 8 kbits/s DTX : Présent Voix : Homme	<ul style="list-style-type: none"> ▪ Dégradations non perçues ▪ Légères atténuations avec des interférences, grésillements, des sons très brouillés et sourds ▪ Fortes atténuations ▪ Atténuations avec de l'écho, des interférences ▪ Coupures 	K ₇₁ K ₇₂ K ₇₃ K ₇₄ K ₇₅
8	Codeur : G729 Débits : 8 kbits/s DTX : Présent Voix : Femme	<ul style="list-style-type: none"> ▪ Dégradations non perçues ▪ Coupures : courtes interruptions et sauts dans la parole ▪ Coupures ▪ Coupures avec du bruit, des crépitements" ▪ Fortes atténuations ▪ Métalliques (faibles) ▪ Bruits supplémentaires : bruits de souffle et grésillements 	K ₈₁ K ₈₂ K ₈₃ K ₈₄ K ₈₅ K ₈₆ K ₈₇
9	Codeur : G723.1 Débits : 5.3 kbits/s DTX : Absent Voix : Femme	<ul style="list-style-type: none"> ▪ Dégradations non perçues ▪ Fortes atténuations, voire coupures ▪ Métalliques, grésillements ▪ Légers grésillements et bips (sur les mots "perturbation" et "gare") ▪ Atténuations et étouffement. 	K ₉₁ K ₉₂ K ₉₃ K ₉₄ K ₉₅
10	Codeur : G723.1 Débits : 5.3 kbits/s DTX : Absent Voix : Homme	<ul style="list-style-type: none"> ▪ Dégradations non perçues ▪ Baisse d'énergie souffle ▪ Baisse d'énergie ▪ Voix déformée, un peu métallique ▪ Métallique, robot, bip. 	K ₁₀₁ K ₁₀₂ K ₁₀₃ K ₁₀₄ K ₁₀₅

A l'observation du tableau précité, on constate que la bibliothèque des défauts permet de mettre en évidence des défauts subjectifs très différents pour des défauts objectifs correspondant à des pertes de paquets
5 remplacés par interpolation puis par du silence, c'est-à-dire pour des défauts objectifs correspondant au même processus de remplacement des paquets perdus par interpolation puis par des trames de silence dépendant notamment du type de codeurs utilisés, du débit de
10 transmission, de la présence ou de l'absence de DTX, du sexe du locuteur et certainement en fonction de la position de la dégradation dans le signal de parole.

Ainsi, la bibliothèque de défauts subjectifs comporte des groupes de défauts perçus catégorisés et
15 qualifiés par une pluralité d'utilisateurs de référence soumis à l'écoute de signaux de parole de test affectés de perte spécifique de paquets.

De préférence, les groupes de défauts, tels que représentés au tableau précité, peuvent en outre être
20 soumis à une procédure de quantification de la gêne perçue permettant d'engendrer une grandeur de quantification de gêne.

En ce qui concerne le module 3 de détection des types de défauts, on indique que le module logiciel
25 précité peut comprendre avantageusement des sous-modules de détection permettant la mise en œuvre du procédé objet de la présente invention, tel que représenté en figure 1 précédemment décrite.

Ainsi, le module 3 de détection des types de
30 défauts assure la détection dans le signal de parole

reconstitué spr des paquets perdus et remplacés par du silence.

Le module précité peut correspondre aux différentes étapes de la figure 1 et comporter un sous-module de calcul de l'énergie moyenne du signal de parole reconstitué dans chaque fenêtre d'analyse, selon l'étape A, un sous-module de détection de chaque fenêtre d'analyse dont l'énergie dans cette fenêtre de rang j est sensiblement nulle, ainsi que représenté à l'étape B de la figure 1.

En outre, le module 3 précité peut comporter un sous-module de détection de la décroissance de l'énergie du signal de parole reconstitué spr précédant le remplacement par des trames de silence. Ce sous-module correspond sensiblement aux étapes C_{11}) et C_{12}) de la figure 1.

Enfin, le module 3 de détection des défauts présents sur le signal de parole reconstitué peut bien entendu comporter, en outre, un sous-module de localisation du défaut dans le corpus du signal de parole, dans un segment de parole ou dans un segment de silence, conformément aux opérations réalisées à l'étape C_{21}) de calcul d'un changement de valeur sur l'enveloppe du signal de parole précité, suivi du test C_{22}) de comparaison de cette valeur de changement à une valeur de seuil, seuil maximum ou seuil minimum, ainsi que décrit précédemment dans la description.

On rappelle que, dans certains cas, le sous-module de calcul d'un changement de valeur sur l'enveloppe à l'étape C_{21}) peut être remplacé par le calcul de la dérivée de l'enveloppe du signal de parole.

On rappelle également que la localisation du défaut dans le corpus du signal de parole, dans un segment de parole ou dans un segment de silence, peut être réalisée à partir de la valeur de la dérivée de l'enveloppe du signal de parole, ainsi que représenté précédemment dans la description en liaison avec la figure 2c.

On rappelle enfin que l'ensemble des étapes de mise en œuvre du procédé objet de l'invention au moyen des modules logiciels précités, et finalement du module 3 de détection des types de défauts, permet de valider la perte de paquets dans des segments de parole à l'étape C₃) et finalement de disposer des variables de type de défauts tdf, lesquelles sont utilisées par le module de calcul et de comparaison vis-à-vis des valeurs de défauts contenus dans la bibliothèque, dans les conditions qui seront explicitées ci-après.

D'une manière générale, on indique que le module de calcul 4, constitué par le module 4a de calcul et de comparaison des types de défaut tdf et des valeurs de défaut de la bibliothèque tdfb, associé bien entendu à l'unité centrale de calcul UC et à l'unité d'affichage UF du micro-ordinateur 4b, permet d'afficher des paramètres de qualité du signal de parole transmis par la mise en œuvre de sous-modules logiciels, lesquels sont illustrés de manière non limitative en figure 3c.

Ainsi que représenté sur la figure précitée, le module 4a de calcul et de comparaison peut comporter avantageusement un sous-module 4a₁ d'identification et de quantification de la gêne des défauts détectés par comparaison.

A titre d'exemple non limitatif, on indique que, par comparaison des types de défauts objectifs tdf et des valeurs de défauts de bibliothèque tdfb, lorsque la comparaison est réussie, on obtient, à partir du tableau précédemment mentionné dans la description, non seulement le type de défauts subjectifs précédemment cité mais également la valeur de gêne relative obtenue à partir de la qualification par les sujets soumis à l'audition des défauts. On rappelle que dans le tableau précité, les valeurs de gêne relative sont notées :

- K_{11} à K_{15} pour le test de défaut numéro 1 ;
- K_{21} à K_{24} pour le test numéro 2 ;
- K_{31} à K_{36} pour le test numéro 3 ;
- K_{41} à K_{45} pour le test numéro 4 ;
- K_{51} à K_{54} pour le test numéro 5 ;
- K_{61} à K_{68} pour le test numéro 6 ;
- K_{71} à K_{75} pour le test numéro 7 ;
- K_{81} à K_{87} pour le test numéro 8 ;
- K_{91} à K_{95} pour le test numéro 9 ;
- K_{101} à K_{105} pour le test numéro 10.

Ces valeurs de gêne relative sont des valeurs qui peuvent être établies par les sujets soumis à l'audition sur une échelle de valeurs spécifique.

Le sous-module 4a₁ peut alors être suivi d'un sous-module 4a₂ permettant de calculer, d'une part, l'occurrence du défaut considéré et, d'autre part, la durée de ce défaut objectif. Le sous-module 4a₂ permet, à partir des valeurs de gêne relative précédemment obtenues, de quantifier la gêne effective à partir de la durée et de l'occurrence des défauts.

Le sous-module 4a₂ peut lui-même être suivi d'un sous-module 4a₃ d'évaluation de la perception des défauts en fonction de leur emplacement.

Alors que l'emplacement du défaut a pu être établi à partir du module 3 précédemment décrit de détection des types de défaut, le module 4a₃ peut consister en un module statistique prenant en compte un effet de rescence mis en évidence à partir d'un ensemble de sujets soumis à une audition des défauts considérés, en fonction de la position de ces défauts dans l'échantillon vocal.

En effet, lors de l'établissement du jugement global, les derniers instants de la séquence sont plus influents, conformément à l'effet de rescence. L'impact des dégradations est ainsi pondéré en fonction de l'emplacement de ces défauts dans la séquence.

D'une manière générale, on rappelle que le modèle statistique correspondant peut être mis en œuvre à partir d'un ensemble d'auditeurs pour des positions de défaut, c'est-à-dire de suppression de paquets remplacés par interpolation puis par du silence intervenant, soit au début, soit au milieu ou encore à la fin des segments de signaux de parole considérés.

Enfin, le sous-module 4a₃ est suivi d'un sous-module 4a₄ d'estimation de la qualité vocale du signal de parole transmis, à partir bien entendu du signal de parole reconstitué spr.

Le sous-module 4a₄ d'estimation de la qualité vocale comporte en fait un logiciel de conversion des paramètres de l'occurrence et de durée des défauts en un paramètre de qualité vocale perçue.

REVENDICATIONS

1. Système de détection non intrusive des défauts d'un signal de parole transmis en téléphonie sur réseau de transmission par paquets, caractérisé en ce qu'il comporte
- 5 au moins :
- des moyens de capture et de reconstitution dudit signal de parole transmis délivrant un signal de parole reconstitué ;
 - des moyens de base de données comportant une

10 bibliothèque de signaux de défauts objectifs, liés aux caractéristiques physiques du signal de parole, et de signaux de défauts subjectifs, liés au typage de la gêne occasionnée à des utilisateurs par lesdits signaux de défauts objectifs, ladite bibliothèque

15 étant construite à partir d'une analyse statistique du type de dégradation ou de défaut présent dans le signal de parole reconstitué ;

 - des moyens de détection de défauts présents sur ledit signal de parole reconstitué et en conséquence sur

20 ledit signal de parole transmis ;

 - des moyens de calcul et d'affichage par comparaison des défauts objectifs présents sur ledit signal de parole reconstitué auxdits signaux de défaut objectifs et/ou signaux de défauts subjectifs de la bibliothèque

25 permettant, à partir d'une analyse statistique du type de dégradation ou de défaut présent dans le signal de parole reconstitué et dans le signal de parole transmis, d'afficher des paramètres de qualité dudit signal de parole transmis.
- 30 2. Système selon la revendication 1, caractérisé en ce que ledit signal de parole transmis en téléphonie

sur réseau de transmission par paquets étant reconstitué à partir d'un réseau local ou d'un réseau longue distance, lesdits moyens de capture et de reconstitution dudit signal de parole comprennent un analyseur de protocoles permettant de capturer et de désencapsuler le signal de parole à partir des paquets transmis, ledit signal de parole capturé et désencapsulé étant mémorisé pour constituer ledit signal de parole reconstitué.

3. Système selon la revendication 1, caractérisé en ce que ledit signal de parole transmis en téléphonie sur réseau de transmission par paquets étant décodé, acheminé et reconstitué à partir du réseau téléphonique commuté, lesdits moyens de capture et de reconstitution du signal de parole comprennent une sonde à haute impédance connectée à une liaison numérique.

4. Système selon l'une des revendications 1 à 3, caractérisé en ce que ladite bibliothèque de défauts subjectifs comporte des groupes de défauts perçus catégorisés et qualifiés par une pluralité d'utilisateurs de référence soumis à l'écoute de signaux de parole de test affectés de pertes spécifiques de paquets.

5. Système selon la revendication 4, caractérisé en ce que les groupes de défauts sont en outre soumis à une procédure de quantification de la gêne perçue, permettant d'engendrer une grandeur de quantification de gêne.

6. Système selon l'une des revendications 1 à 5, caractérisé en ce que lesdits moyens de détection de défauts présents sur ledit signal de parole reconstitué comportent au moins un module de détection, dans le signal

de parole reconstitué, des paquets perdus remplacés par du silence.

5 7. Système selon l'une des revendications 1 à 6, caractérisé en ce que lesdits moyens de détection de défauts présents sur ledit signal de parole reconstitué comportent en outre un module de détection de la décroissance de l'énergie du signal de parole reconstitué, précédant le remplacement par des trames de silence.

10 8. Système selon l'une des revendications 1 à 7, caractérisé en ce que lesdits moyens de détection de défauts présents sur ledit signal de parole reconstitué comportent en outre un module de localisation du défaut dans le corpus du signal de parole, dans un segment de parole ou dans un segment de silence.

15 9. Système selon l'une des revendications 1 à 8, caractérisé en ce que lesdits moyens de calcul et d'affichage permettant d'afficher des paramètres de qualité du signal de parole transmis comprennent au moins, pour tout type de défaut détecté :

- 20 - des moyens de calcul de l'occurrence du défaut ;
- des moyens de calcul de la durée du défaut ;
- des moyens de conversion des paramètres d'occurrence et de durée des défauts en un paramètre de qualité vocale perçue du signal de parole.

25 10. Système selon la revendication 9, caractérisé en ce que lesdits moyens de conversion des paramètres d'occurrence de durée des défauts en un paramètre de qualité vocale perçue du signal de parole comportent des moyens de calcul d'une fonction de pondération de
30 l'importance subjective des défauts objectifs détectés,

suivant leur emplacement temporel vis-à-vis du segment de parole.

11. Procédé de détection des défauts objectifs d'un signal de parole transmis par paquets en téléphonie sur réseau de transmission par paquets, caractérisé en ce qu'il consiste, à partir d'échantillons d'un signal de parole reconstitué représentatif de ce signal de parole transmis et analysés sur au moins une fenêtre d'analyse :

- à calculer l'énergie moyenne du signal de parole reconstitué dans cette fenêtre d'analyse ;
- à rechercher, pour une succession de fenêtres d'analyse, les segments du signal de parole reconstitué à énergie moyenne nulle ; et, sur existence d'au moins une fenêtre d'analyse d'énergie moyenne nulle,
- à valider la perte de paquets avec substitution par des trames de silence conditionnellement à l'existence d'une substitution de trame.

12. Procédé selon la revendication 11, caractérisé en ce que l'étape consistant à valider la perte de paquets IP consiste :

- à calculer la pente de décroissance de l'énergie d'excitation de trames antérieures successives ;
- à valider la perte de paquets en fonction de l'existence d'un défaut objectif dans un segment d'activité vocale.

13. Procédé selon la revendication 11, caractérisé en ce que l'étape consistant à valider la perte de paquets en fonction de l'existence d'un défaut objectif dans un segment d'activité vocale consiste :

- à calculer une valeur de changement de valeur de l'enveloppe du signal de parole reconstitué ;
- à comparer ladite valeur de changement à au moins une valeur de seuil, une réponse positive à ladite comparaison révélant l'existence d'une perte de paquets dans un segment d'activité vocale engendrant un défaut subjectif et une réponse négative à ladite comparaison révélant l'absence de défaut subjectif.

1/6

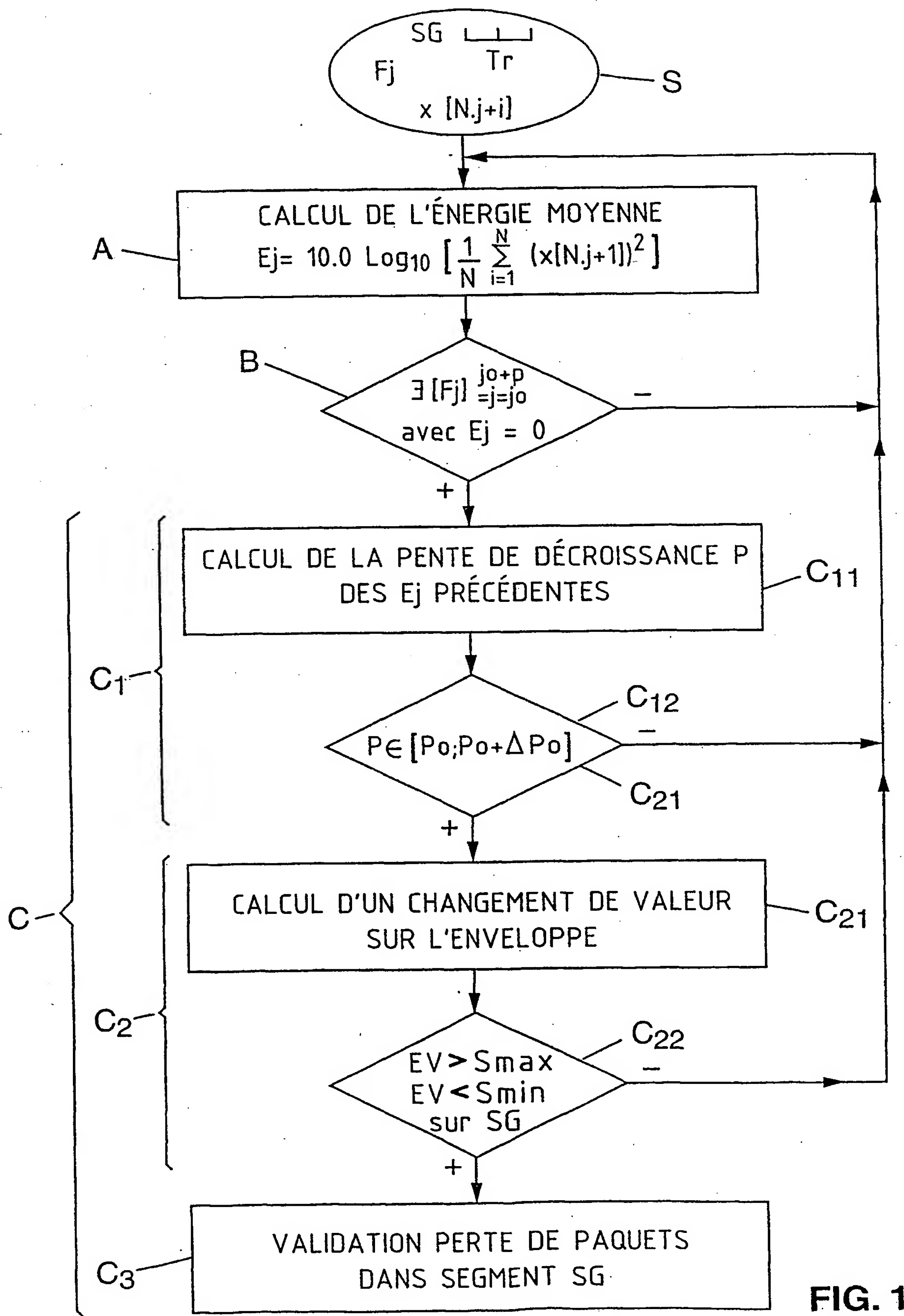


FIG. 1

2/6

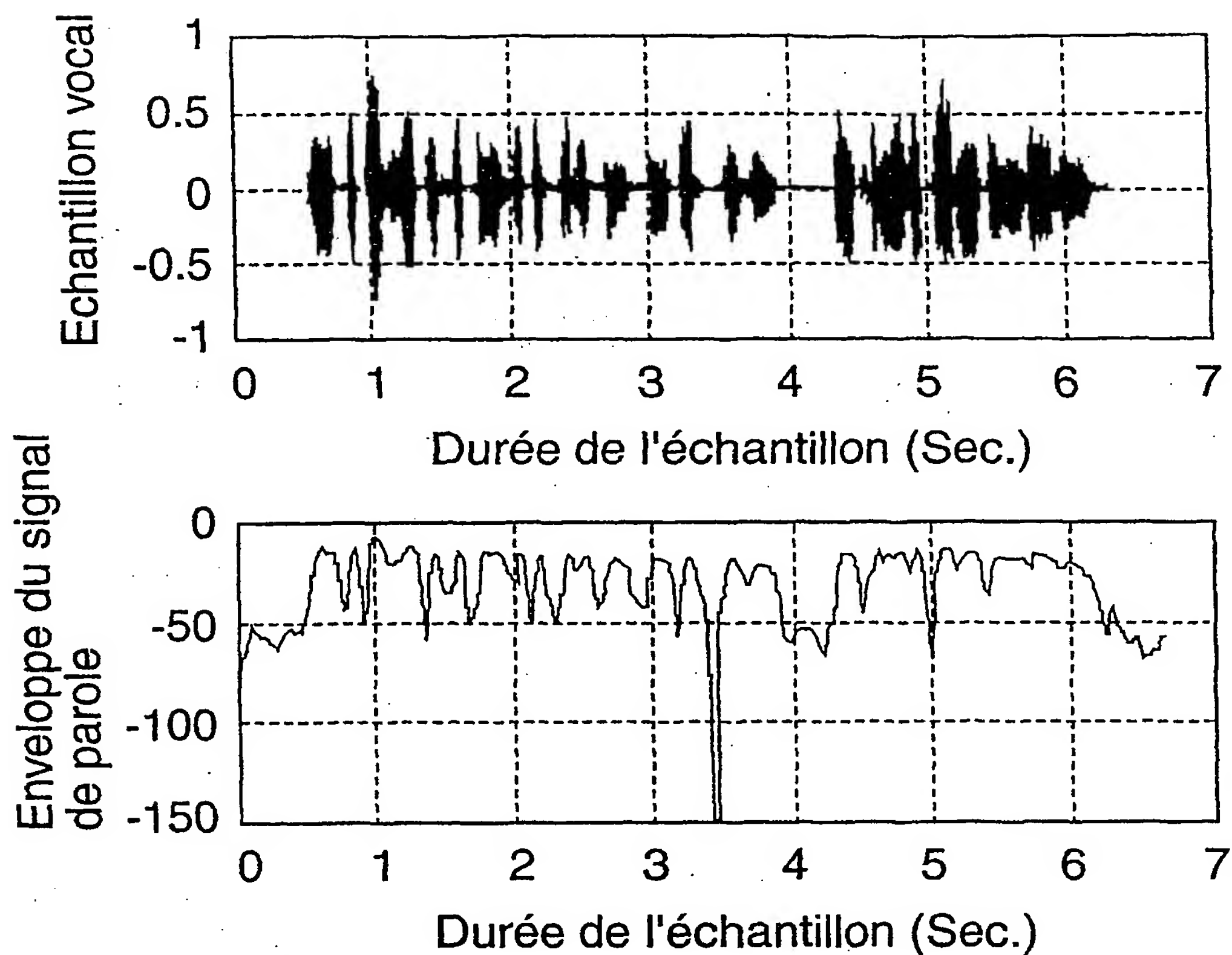


FIG. 2a : Echantillon de parole (après décodage G723) affecté d'une perte de 5 paquets.

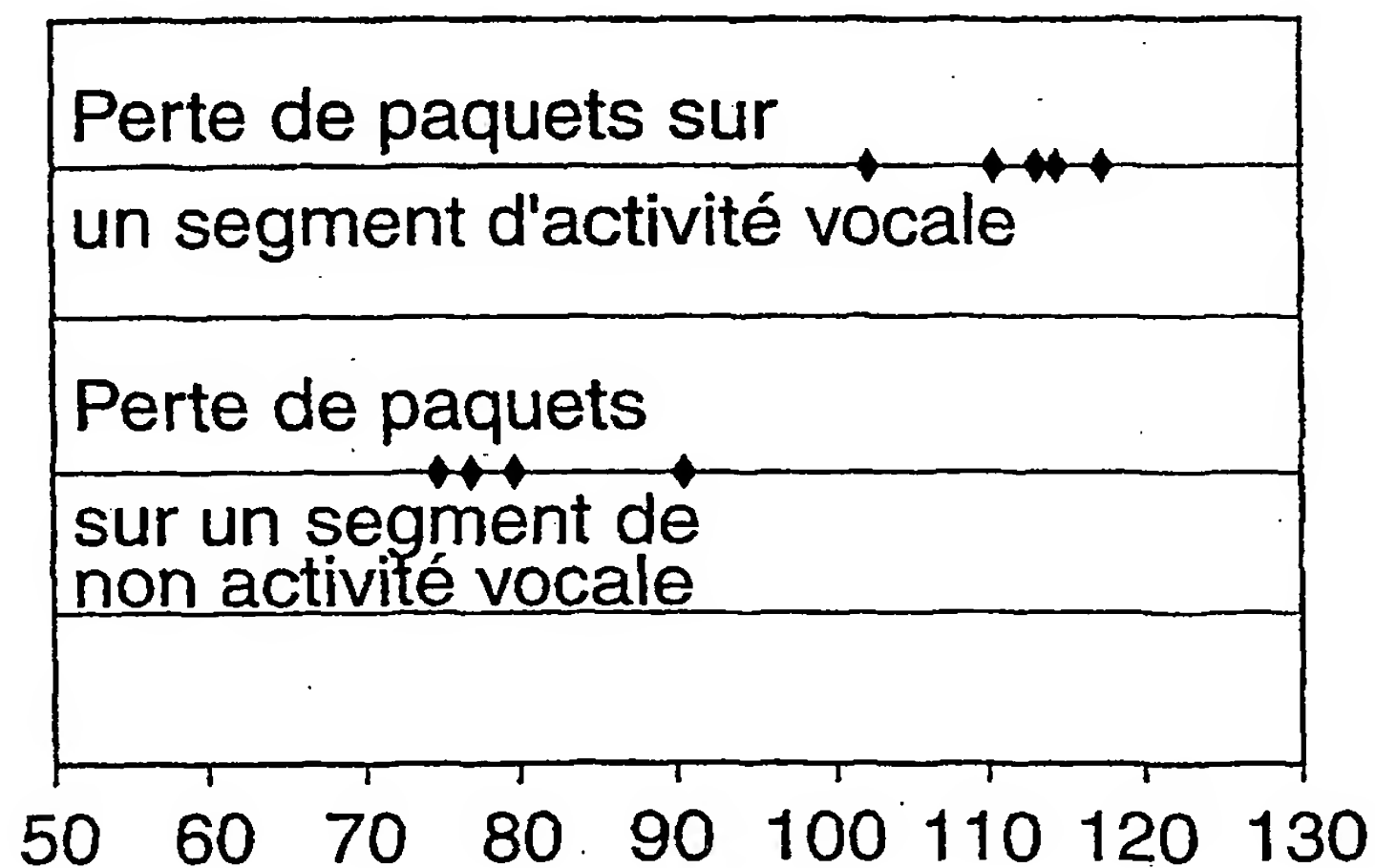


FIG. 2c

3/6

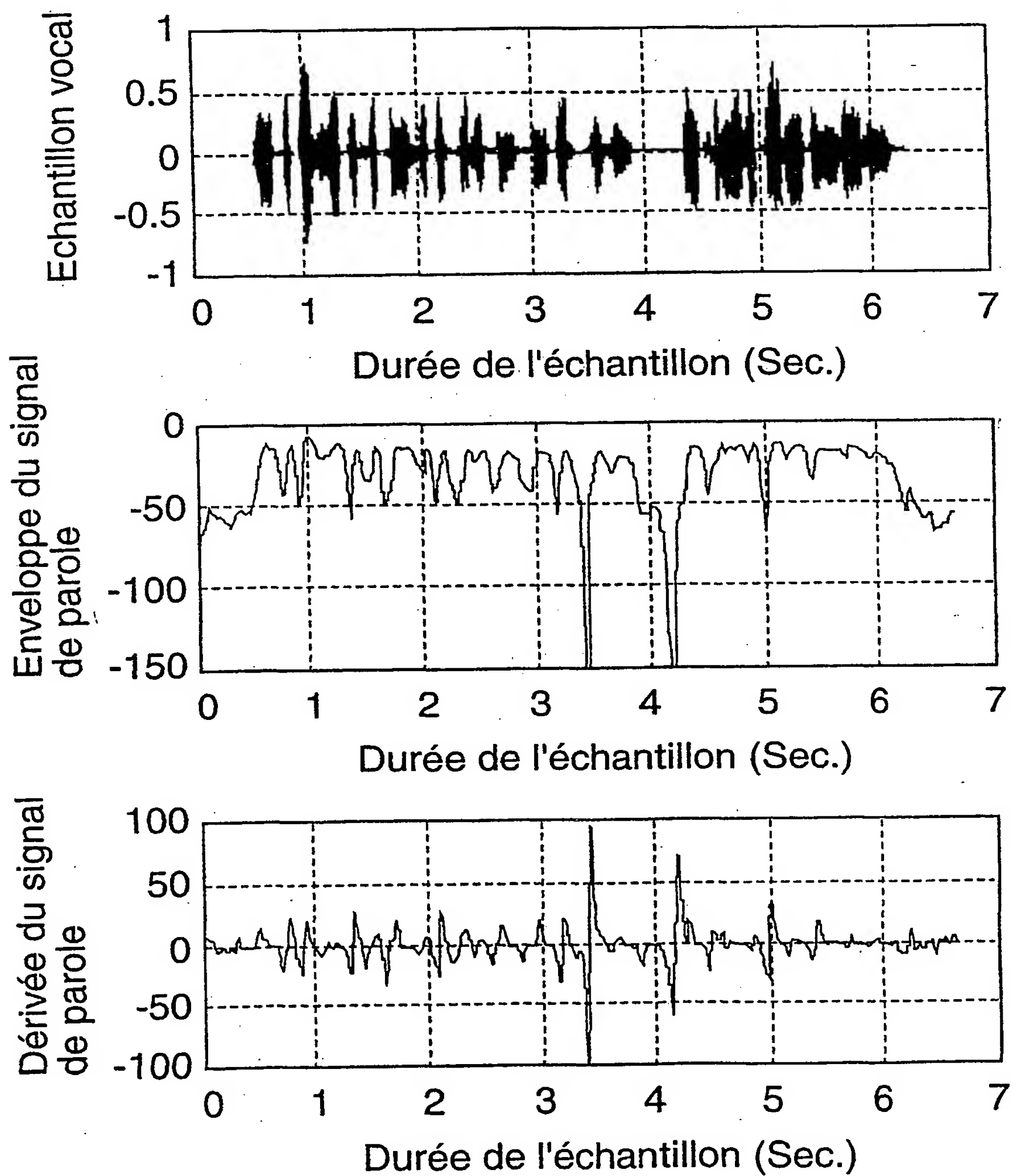


FIG. 2b

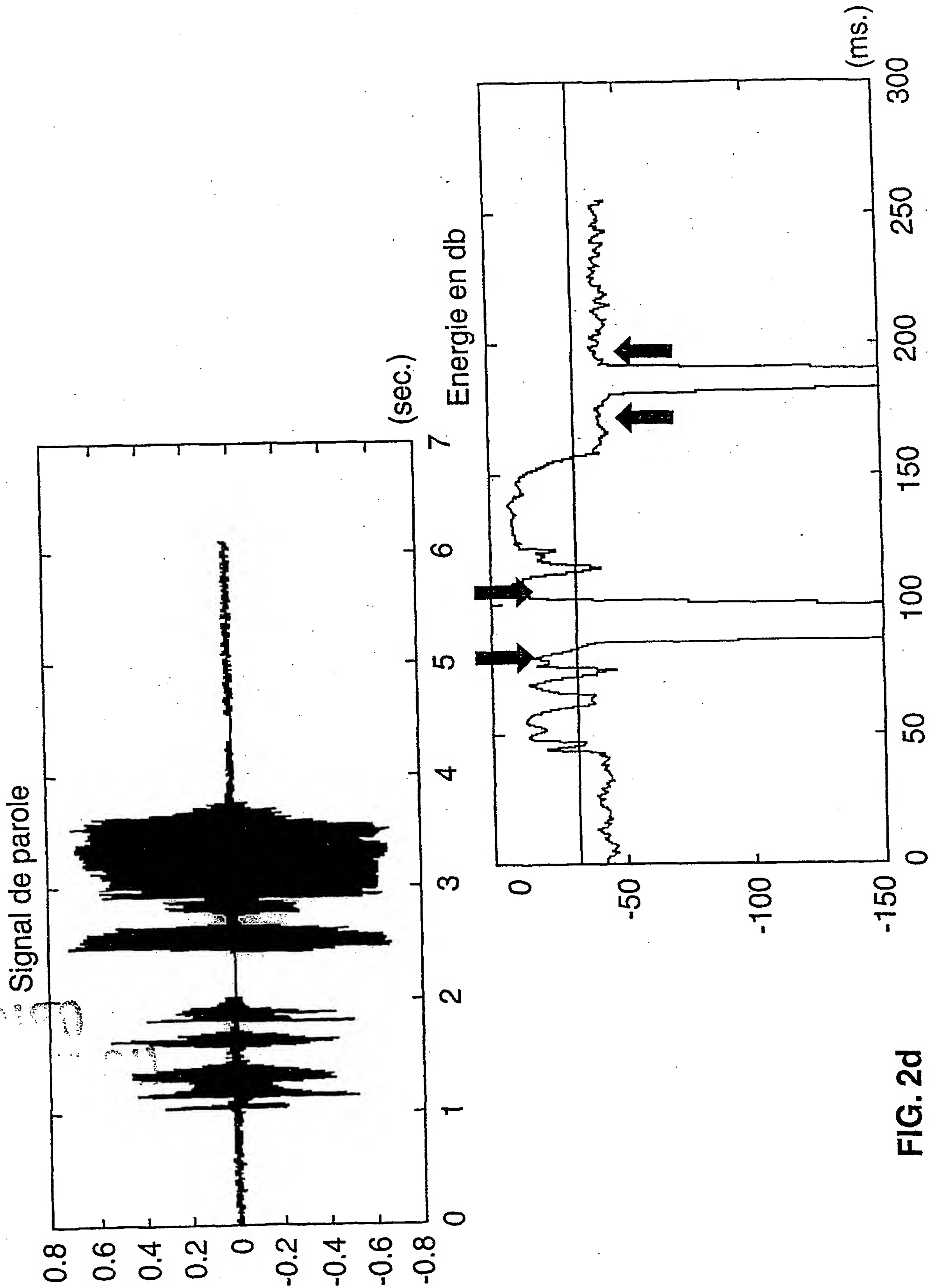
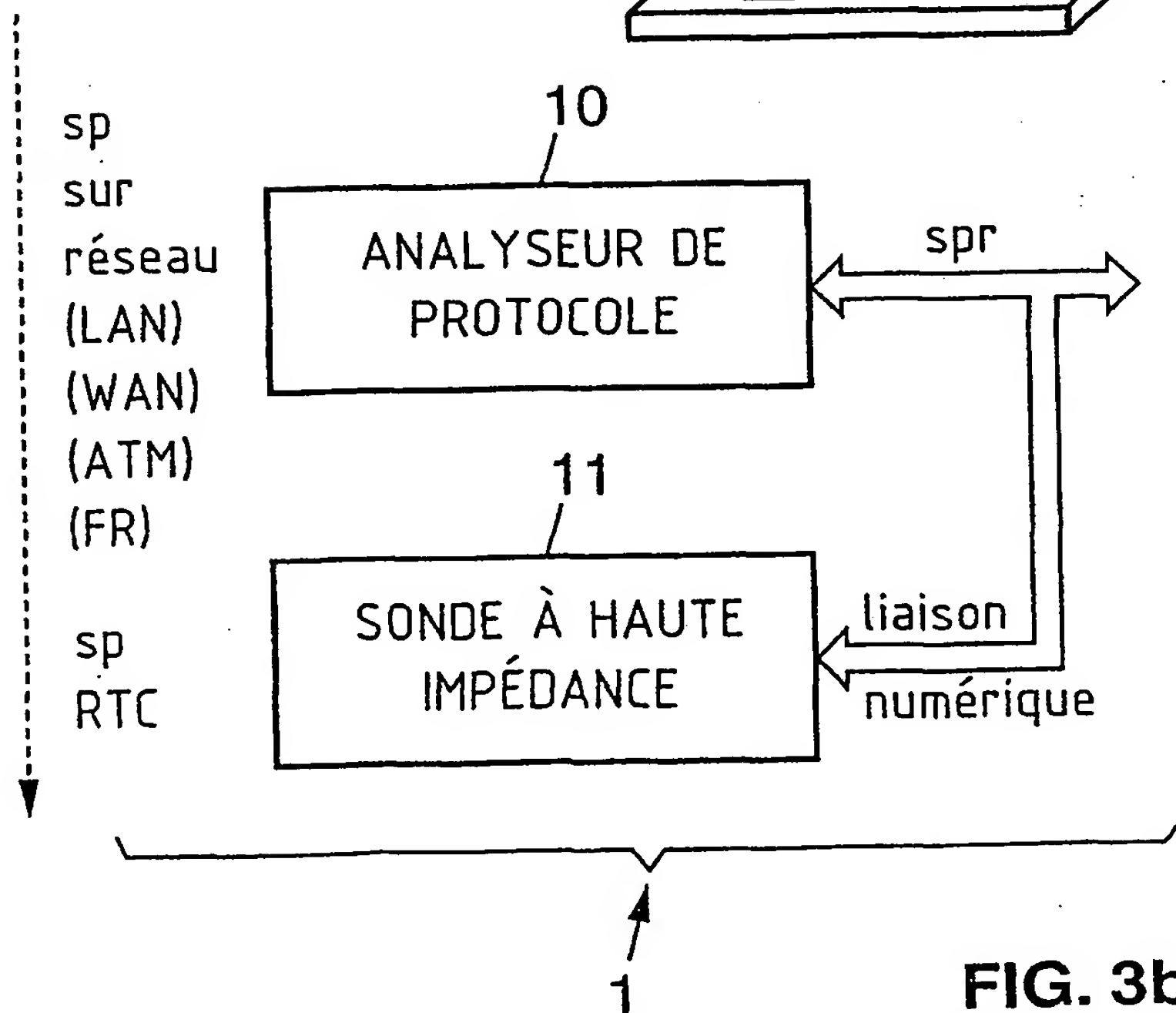
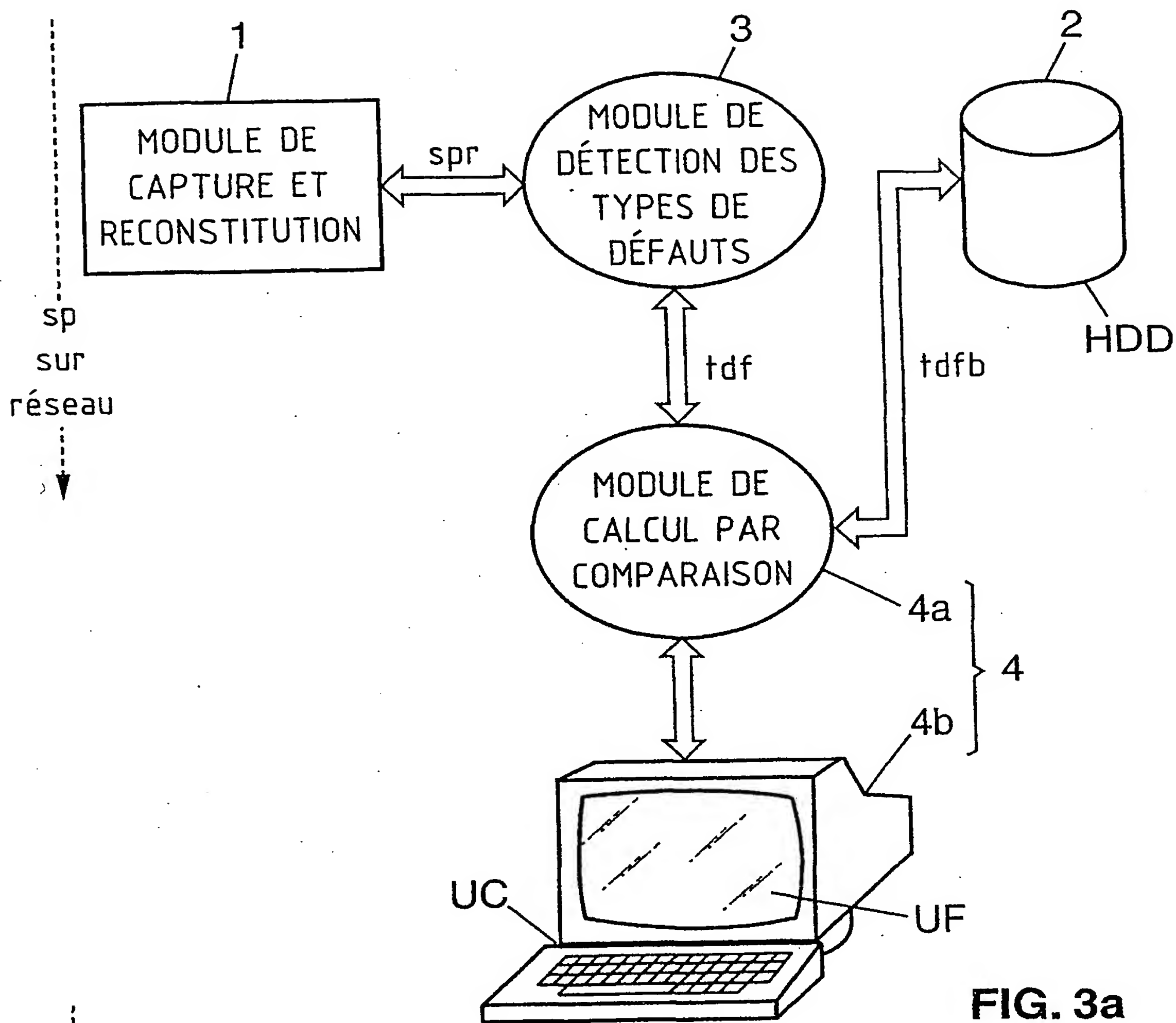


FIG. 2d

5/6



ORIGINAL
NO MARKING

6/6

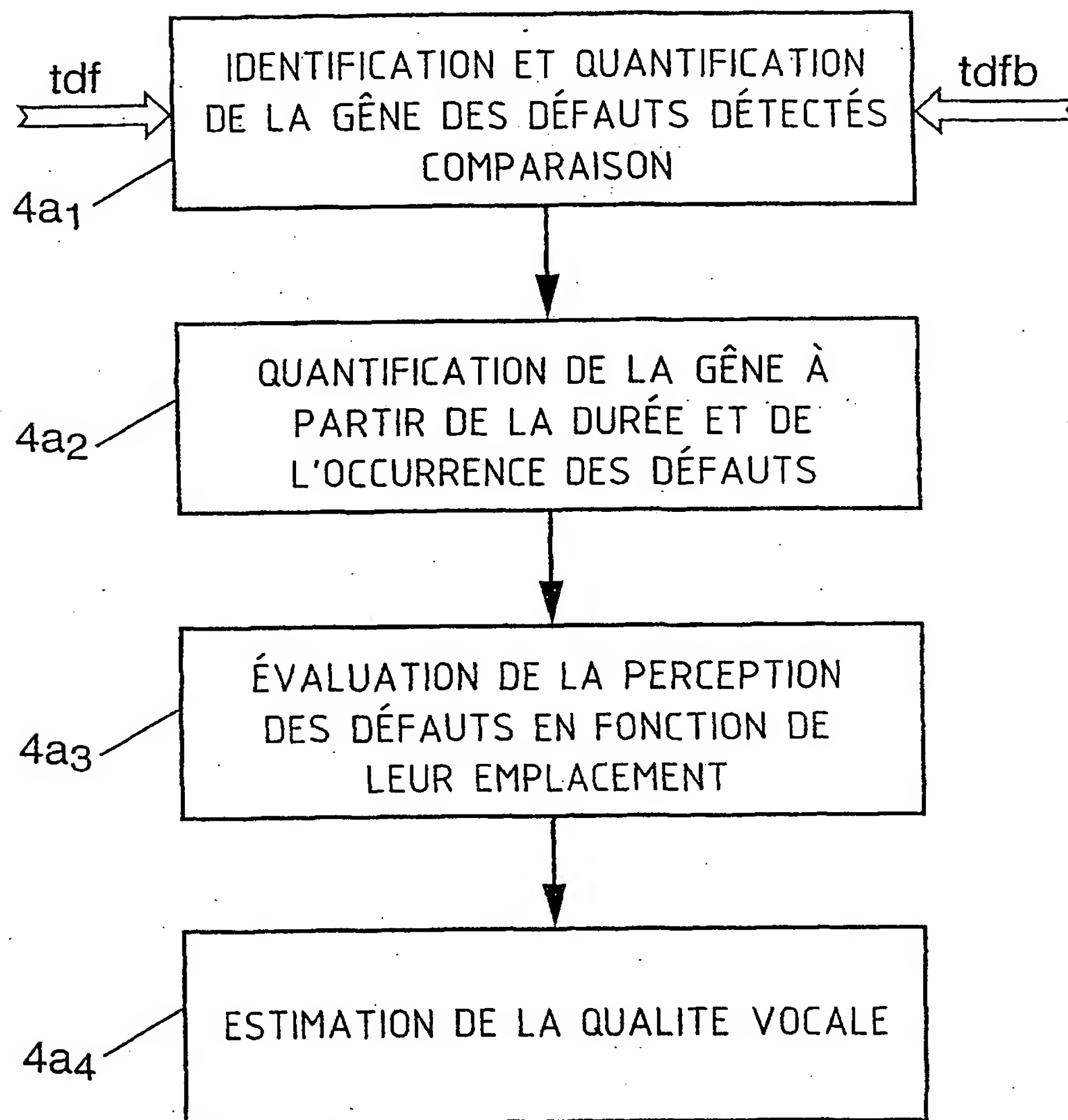


FIG. 3c

INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 01/03671

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G10L19/00 H04L1/20 H04L12/56

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G10L H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

INSPEC, EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BERNEX E.: "Caractérisation et mesure des dégradations de la qualité vocale téléphonique sur des réseaux à qualité vocale non garantie" RENCONTRES JEUNES CHERCHEURS EN PAROLE, 18 - 19 November 1999, XP002171229 Avignon, France the whole document --- -/-	1,2,4-6, 8-12

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *8* document member of the same patent family

Date of the actual completion of the international search

5 February 2002

Date of mailing of the international search report

12/02/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Quélavoine, R

INTERNATIONAL SEARCH REPORT

International Application No
PCT/FR 01/03671

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>PRUTHI P ET AL: "Application level performance of multimedia services" INTERNET II: QUALITY OF SERVICE AND FUTURE DIRECTIONS, BOSTON, MA, USA, 20-21 SEPT. 1999, vol. 3842, pages 154-164, XP001012099 Proceedings of the SPIE - The International Society for Optical Engineering, 1999, SPIE-Int. Soc. Opt. Eng, USA ISSN: 0277-786X abstract page 9, paragraph 2 -page 10, paragraph 2</p>	1,2,4-6, 11
A	<p>US 6 119 083 A (HOLLIER ET AL) 12 September 2000 (2000-09-12) abstract</p>	1,4,5
A	<p>GHITA B V ET AL: "IP network performance monitoring of voice flows for IP telephony" INC 2000. PROCEEDINGS OF THE 2ND INTERNATIONAL NETWORK CONFERENCE, PROCEEDINGS OF INC 2000. 2ND INTERNATIONAL NETWORK CONFERENCE, PLYMOUTH, UK, 3-6 JULY 2000, pages 145-155, XP001011498 2000, Plymouth, UK, Univ. Plymouth, UK ISBN: 1-84102-066-4 abstract</p>	1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/FR 01/03671

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6119083	A	12-09-2000	
		AU 711615 B2	14-10-1999
		AU 1553197 A	16-09-1997
		AU 694932 B2	06-08-1998
		AU 6623296 A	26-02-1997
		CA 2225407 A1	13-02-1997
		CA 2237814 A1	04-09-1997
		CN 1192309 A	02-09-1998
		EP 0840975 A1	13-05-1998
		EP 0879527 A1	25-11-1998
		WO 9705730 A1	13-02-1997
		WO 9732428 A1	04-09-1997
		JP 11510334 T	07-09-1999
		JP 2000506327 T	23-05-2000
		NO 980331 A	26-01-1998
		NO 983959 A	28-08-1998
		NZ 313705 A	25-11-1998
		US 6035270 A	07-03-2000
		US 5799133 A	25-08-1998

RAPPORT DE RECHERCHE INTERNATIONALE

Der Internationale No
PCT/FR 01/03671

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
CIB 7 G10L19/00 H04L1/20 H04L12/56

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)
CIB 7 G10L H04L

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)
INSPEC, EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	BERNEX E.: "Caractérisation et mesure des dégradations de la qualité vocale téléphonique sur des réseaux à qualité vocale non garantie" RENCONTRES JEUNES CHERCHEURS EN PAROLE, 18 - 19 novembre 1999, XP002171229 Avignon, France le document en entier --- -/--	1,2,4-6, 8-12

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

* Catégories spéciales de documents cités:

- *A* document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- *E* document antérieur, mais publié à la date de dépôt international ou après cette date
- *L* document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- *O* document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- *P* document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

- *T* document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention
- *X* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément
- *Y* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier
- *Z* document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

5 février 2002

Date d'expédition du présent rapport de recherche internationale

12/02/2002

Nom et adresse postale de l'administration chargée de la recherche internationale
Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Quélavoine, R

RAPPORT DE RECHERCHE INTERNATIONALE

Den internationale No

PC171R 01/03671

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>PRUTHI P ET AL: "Application level performance of multimedia services"</p> <p>INTERNET II: QUALITY OF SERVICE AND FUTURE DIRECTIONS, BOSTON, MA, USA, 20-21. SEPT. 1999,</p> <p>vol. 3842, pages 154-164, XP001012099</p> <p>Proceedings of the SPIE - The International Society for Optical Engineering, 1999, SPIE-Int. Soc. Opt. Eng, USA</p> <p>ISSN: 0277-786X</p> <p>abrégé</p> <p>page 9, alinéa 2 -page 10, alinéa 2</p>	1,2,4-6, 11
A	<p>US 6 119 083 A (HOLLIER ET AL)</p> <p>12 septembre 2000 (2000-09-12)</p> <p>abrégé</p>	1,4,5
A	<p>GHITA B V ET AL: "IP network performance monitoring of voice flows for IP telephony"</p> <p>INC 2000. PROCEEDINGS OF THE 2ND INTERNATIONAL NETWORK CONFERENCE, PROCEEDINGS OF INC 2000: 2ND INTERNATIONAL NETWORK CONFERENCE, PLYMOUTH, UK, 3-6 JULY 2000,</p> <p>pages 145-155, XP001011498</p> <p>2000, Plymouth, UK, Univ. Plymouth, UK</p> <p>ISBN: 1-84102-066-4</p> <p>abrégé</p>	1

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Der Internationale No

PCT/FR 01/03671

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 6119083 A	12-09-2000	AU 711615 B2	14-10-1999
		AU 1553197 A	16-09-1997
		AU 694932 B2	06-08-1998
		AU 6623296 A	26-02-1997
		CA 2225407 A1	13-02-1997
		CA 2237814 A1	04-09-1997
		CN 1192309 A	02-09-1998
		EP 0840975 A1	13-05-1998
		EP 0879527 A1	25-11-1998
		WO 9705730 A1	13-02-1997
		WO 9732428 A1	04-09-1997
		JP 11510334 T	07-09-1999
		JP 2000506327 T	23-05-2000
		NO 980331 A	26-01-1998
		NO 983959 A	28-08-1998
		NZ 313705 A	25-11-1998
		US 6035270 A	07-03-2000
		US 5799133 A	25-08-1998